Online Supplementary S1

Proof for Proposition 1

Proof for Statement 1

In order to prove Eq (3), we first prove the following convergence relation：

$$\forall\delta > 0 \lim_{n\to\infty} \Pr\left(\|\hat{F}_{N_n} - F\|_{KS} < \delta\right) = 1 \tag{A1}$$

For any $\delta > 0$ and any positive integer n, define a set $F_{n,\delta} = \{N \subset M_n\colon \|\hat{F}_N - F\|_{KS} \geq \delta\}$. Obviously, Eq (A1) holds if for anyδ, with $n \to \infty$ we have $\Pr\left(\cup_{N\in F_{n,\delta}}\{N = N_n\}\right) \to 0$, where $\{N = N_n\}$ denotes the set of all random states such that the optimal market share N_n in period n is equal to the given consumer base N at these states. It is easy to see that for a given δ, there exists a ratio $0 < p_\delta < 1$, such that $|N| \leq p_\delta |M_n|$ holds for all n.

For each $N \in F_{n,\delta}, \{N = N_n\} = \cap_{l\in\frac{M_n}{N}}\{l \notin N_n\}$, i.e., the event "$N$ equals the customer base N_n under some optimal interest rate" is equivalent to the event "Every consumer not in N is not in the optimal customer base N_n". For any $l \notin N$, that $l \notin N_n$ if and only if for all $n_l \leq m \leq n$, $R(y_l, \omega_l) \leq r_m^*$ always holds. Based on Eq (1), $R(y_l, \omega_l) \leq r_m^*$ holds iff $\left(1 + R(y_l, \omega_l)\right)\left(1 - p(y, \omega'_{m,l})\right) - p(y_l, \omega'_{m,l}) < 0$ holds, where $\omega'_{m,l}$ is the inferred value of the true risk state ω_l for consumer l in period m. According to assumptions 2 and 3, plus $\delta > 0$, $\Pr\left(\left(1 + R(y_l, \omega_l)\right)\left(1 - p(y, \omega'_{m,l})\right) - p(y_l, \omega'_{m,l}) < 0\right) \in \left[\tau_\delta, \overline{\tau}_\delta\right]$, where the upper and lower bounds $\tau_\delta > 0, \overline{\tau}_\delta < 1$, we have $\Pr(l \notin N_n) \in \left[\tau_\delta^{n-n_l+1}, \overline{\tau}_\delta^{n-n_l+1}\right]$, and the independence of intertemporal and cross-sectional observation errors $\Pr(N = N_n) \in \left[\prod_{l\in\frac{M_n}{N}}\tau_\delta^{n-n_l+1}, \prod_{l\in\frac{M_n}{N}}\overline{\tau}_\delta^{n-n_l+1}\right]$. $\left|\frac{M_n}{N}\right| > (1 - p_\delta)|M_n|$ and assumption 1 indicate that $\Pr(N = N_n)$ is a higher order infinitesimal than$\tau_\delta^{(1-p_\delta)^2|M_n|^2-(1-p_\delta)|M_n|}$, where$\tau_\delta^2 \in \left[\tau_\delta, \overline{\tau}_\delta\right]$.

On this basis, for a positive integer $k \leq p|M_n|$, $F_{n,\delta,k} = \left\{N \in F_{n,\delta}\colon |N| = k\right\}$, then $\left|F_{n,\delta,k}\right| = C(|M_n|, p_k|M_n|) = \frac{|M_n|!}{((1-p_k)|M_n|)!(p_k|M_n|)!}$ is an infinite amount of the same order as or lower than $\tau^{-\alpha|M_n|}$, where $p_k = \frac{k}{|M_n|}$, and α and τ are non-zero positive constants with $\alpha < 1$. Thus, $\left|F_{n,\delta}\right|$ is an infinity of no higher order than $p_\delta|M_n|\tau^{-\alpha|M_n|}$, which implies that

$$\Pr\left(\bigcup_{N\in F_{n,\delta}}\{N = N_n\}\right) \leq C\tau_\delta^{(1-p_\delta)^2|M_n|^2-(1-p_\delta)|M_n|-\alpha'|M_n|-\tau'\log|M_n|}$$

where C, α' and τ' are all nonzero positive constants. The inequality above proves Eq (A1).

Given Eq (A1), The convergence relation (3) is equivalent to $\lim_{n\to\infty}\Pr(\cup_{N\in F'_{n,\delta}}\{N = N_n\}) = 0$ holds for any $\delta > 0$, where $F'_{n,\delta} = \left\{N \subset M_n: \frac{|N|}{|M_n|} \le \delta\right\}$. From the above proof for Eq. (A1), it follows that for $\delta > 0$, this condition holds only if the following condition holds.

Condition (*): There exist $\tau < 1$ and $p > 0$, such that for a sufficient large n and for any $N \in F'_{n,\delta}$, all $l \in \frac{M_n}{N}$ that satisfy the conditions $\Pr(l \notin N_m) < \tau$, $n_l \le m \le n$, and $n_l = \min\{m \le n: l \in M_m\}$ are in the set $\frac{M_n}{N}$ in a proportion not less than p.

According to Eq (A1), given $\delta > 0$, we can always choose certain δ_1, ϵ_1 and n_1, such that for any $n > n_1$, all random states can be divided into 2 sets $A_{1,n} = \{\|\hat{F}_{N_n} - F\|_{KS} < \delta_1\}$ and $A_{2,n} = \{\|\hat{F}_{N_n} - F\|_{KS} \ge \delta_1\}$. Eq (A1) implies $\Pr(A_{2,n}) \to 0$, leaving only the former. Since $\|\hat{F}_{N_n} - F\|_{KS} < \delta_1$ holds on $A_{1,n}$, Thus for all consumers l who satisfy $(1 + R(y_l,\omega_l))(1 - p(y_l,\omega_l)) - p(y_l,\omega_l)$, the conditional probability that they are not part of the institutional clientele on $A_{1,n}$, $\Pr(l \notin N_m|A_{1,n})$, must be consistently smaller than some positive constant $\tau < 1$.Meanwhile, $\|\hat{F}_{\frac{M_n}{N_n}} - F\|_{KS} < \delta_1$ holds obviously since $\|\hat{F}_{N_n} - F\|_{KS} < \delta_1$, therefore the proportion of consumers I of $(1 + R(y_l,\omega_l))(1 - p(y_l,\omega_l)) - p(y_l,\omega_l)$ in $\frac{M_n}{N_n}$ is not less than $p = \int_\Omega I\Big((1 + R(y,\omega))(1 - p(y,\omega_l)) - p(y,\omega) > 0\Big)\,dF - \delta_1 > 0$. Then Condition (*) holds on $A_{1,n}$.

Proof for Statement 2

We first prove the following convergence relation holds for any $\delta, \epsilon > 0$ and any consumer $l \in \cup_{n>0} M_n$:

$$\lim_{n\to\infty}\Pr(\{r^*_{n,l} \ge \bar{r}^*_{2,l} + \epsilon\} \mid \min_k\|\hat{F}_{N_{n,k}} - F\|_{KS} \ge \delta) = 0 \tag{A2}$$

Note that Eq (A2) implies that for any consumer $l \in \cup_{n>0} M_n$, the following holds

$$\begin{aligned}\Pr\Big(\bigcup_{m>0}\bigcap_{n>n_l}&\left\{r^*_{n,l} \ge \bar{r}^*_2 + \frac{1}{m}\right\} \mid \min_k\|\hat{F}_{N_{n,k}} - F\|_{KS} \ge \delta\Big)\\ &= \Pr\left(\bigcap_{n>n_l}\{l \notin N_{n,k_1}\} \mid \min_k\|\hat{F}_{N_{n,k}} - F\|_{KS} \ge \delta\right) = 0\end{aligned}$$

Thus

$$\begin{aligned}
&\lim_{n\to\infty} \quad \Pr\,(N_n = M_n \mid \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta) \\
&= \quad 1 - \Pr\left(\bigcup_{l\in \cup_{s>0} M_s} \bigcap_{n>n_l} \{\, l \notin N_{n,k_1} \right\} \mid \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta \Big) = 0
\end{aligned}$$

On the other hand, $N_n = M_n$ implies $\lim_{n\to\infty} \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} = 0$, thus we have

$$\begin{aligned}
0 = & \lim_{n\to\infty} \Pr\,(N_n = M_n, \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta) \\
= & \lim_{n\to\infty} \Pr\,(N_n = M_n, \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta) \Pr\,(\min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta) \\
= & \lim_{n\to\infty} \Pr\,(\min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta)
\end{aligned}$$

ie.e, Eq(4) holds, which leaves us the task of proving Eq (A2).

From the definitions of $\overline{r}_2^*$ and r_n^*, we know that for any $\varepsilon > 0$ and any consumer l, there exists infinite number of n such that $r_{n,k_2,l}^* < \overline{r}_{2,l}^* + \frac{\varepsilon}{2}$, Denote the set consisting of all such n's as N^l. Without loss of generality, we assume that for all $n \in N^l$, we have $\min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta$.

Note thatbased on assumptions 2 and 3, we have

$$Pr\left(r_{n,l}^* \geq r_{n,k_2,l}^* + \varepsilon/2 \mid \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta\right) \leq \tau_{\delta,\varepsilon} < 1.$$

From the independence of observation errors, it follows that for any given positive integer m, we have

$$\begin{aligned}
& Pr\left(r_{m,l}^* \geq \overline{r}_{2,l}^* + \varepsilon \mid \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta\right) \\
\leq \;& Pr\left(\bigcap_{n\in N^l, n\leq m} r_{n,l}^* \geq r_{n,k_2,l}^* + \varepsilon/2 \mid \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta\right) \\
= \;& \prod_{n\in N^l, n\leq m} Pr\left(r_{n,l}^* \geq r_{n,k_2,l}^* + \varepsilon/2 \mid \min_k \|\hat{F}_{N_{n,k}} - F\|_{KS} \geq \delta\right) \\
\leq \;& \prod_{n\in N^l, n\leq m} \tau_{\delta,\varepsilon}
\end{aligned}$$

As N^l is an infinite set, the above expression converges to 0. Eq (A2) is proven.

Proof for Statement 3

Next, we prove Eqs (5) and (6). Since Eq (6) is a special case of (5) under $\overline{r}_{i+}^* = r_0$, we only need to prove Eq (5), which can be divided into 2 parts:

$$\lim_{n\to\infty} Pr\left(lr_n^* < r_0 - \varepsilon \mid \|\hat{F}_{N_{n,k_{K,n}}} - F\|_{KS} > 0\right) = 0, \forall \varepsilon > 0, \qquad \text{(A3)}$$

$$\lim_{n\to\infty} Pr\left(lr_n^* > \max(\overline{r}_{i+}^*, r^*) + \varepsilon \mid \|\hat{F}_{N_{n,k_{K,n}}} - F\|_{KS} > 0\right) = 0,\ \forall \varepsilon > 0. \tag{A4}$$

And note that using the same method as in the proof of Eq (A2) it follows that Eq (A4) clearly holds if $\overline{r}_{i+}^* \geq r^*$, and Eq (A4) clearly holds if $\overline{r}_{i+}^* \leq r^*$. With respect to Eq (A3), r_0, by definition, denotes the lowest interest rate quoted by the online lender, and a lower interest rate quote will result in a negative marginal expected return with probability 1, so equation (A3) holds. □